A NEW DESCRIPTOR BASED ON 2D DCT FOR IMAGE RETRIEVAL

Cong BAI, Kidiyo KPALMA and Joseph RONSIN

Université Européenne de Bretagne, France
INSA de Rennes, IETR, UMR 6164, F-35708, RENNES
{cong.bai | kidiyo.kpalma | joseph.ronsin}@insa-rennes.fr

Abstract: Content-based image retrieval relies on feature comparison between images. So the selection of feature vector is important. As many images are compressed by transforms, constructing the feature vector directly in transform domain is a very popular topic. We propose a new feature vector in DCT domain. Our method selects part of DCT coefficients inside each block to construct AC-Pattern and use DC coefficients between neighboring blocks to construct DC-Pattern. Two histograms are formed and parts of them are used to build a descriptor vector integrating features to do image retrieval. Experiments are done both on face image databases and texture image database. Compared to other methods, results show that we can get better performance on both face and texture database by using the proposed method.

1   INTRODUCTION

The rapid increase in the digital image collections gives more and more information. However, difficulty for an efficient use of this information is growing, unless we can browse, search and retrieve it easily. Content-based image retrieval (CBIR) has been an active research field in pattern recognition and computer vision for decades. As much of images are stored in compressed format with different kinds of transforms, image retrieval in transform domain have been widely developed in many researches.

Discrete Cosine Transform (DCT) is used in JPEG compression standard. In another aspect, DCT is also used as a efficient tool to extract features in image retrieval. Consequently in last decades, several researches appeared in DCT-based image retrieval. Composing coefficients for a feature vector that represents the image leads to different solutions. In (Tsai et al., 2006), the upper left DCT coefficients are categorized into four groups, one is DC coefficients and other three includes the coefficients which have vertical, horizontal and diagonal information. And these four groups compose the feature vectors. In (Zhong and Defée, 2005), 2 DCT patterns are generated from the DCT blocks and then their histograms are constructed and combined to do

retrieval. In (Bai et al., 2011), an improved histogram descriptor is obtained by using zig-zag scan and observing adjacent patterns on coefficients.

The wide use of DCT in image compression and image retrieval comes from its capability to compact the energy. It means that much of the energy lies in low frequency coefficients, so that high frequency can be discarded without visible distortion. In other words, only a reduced part of DCT coefficients can efficiently represent the image contents. In comparison to the use of all of the coefficients, this reduces the complexities and redundancies of the features vectors applied for image retrieval. Our method proposed in this paper is inspired from this consideration and previous works (Tsai et al., 2006) (Zhong and Defée, 2005) (Bai et al., 2011).

In this paper, we present a simple but effective approach to construct a descriptor from DCT coefficients for image retrieval. Experimental results show that our method can apply both on face database and texture database, corresponding to different structure of image contents and moreover can achieve better performance than many classical methods and state-of-art approach.

The rest of the paper is organized as follows. Constructing patterns and descriptor vectors are described in section 2. Section 3 presents the

analysis of experimental results both on face and texture databases and finally a conclusion is given in section 4.

2 DESCRIPTION OF METHOD

2.1. General description

In this study, we use 4×4 block DCT transform. So we get 1 DC coefficient and 15 AC coefficients for each block. For each block, we select 9 AC coefficients to construct AC-Pattern, and use DC coefficients of the block itself and DC coefficients of its 8 neighboring blocks to build DC-Pattern. We generate the histogram of pattern as number of appearance of patterns in the image. Finally, we use the concatenation of AC-Pattern histogram and DC-Pattern histogram as the descriptor of the image to do retrieval.

2.2. Pre-Processing

We adopt the luminance normalization method presented in (Zhong and Defée, 2005) as pre-processing steps to eliminate the effect of luminance variations. From these pre-processed coefficients, histograms will be observed considering the occurrence of their contents.

2.3. AC-Pattern and its histogram

As mentioned before, subsets of coefficients can represent the image content. So we will select at most 9 coefficients in each block to construct the AC-Pattern. This selection gathers these 9 coefficients into 3 groups: horizontal, vertical and diagonal, as shown in Fig.1.
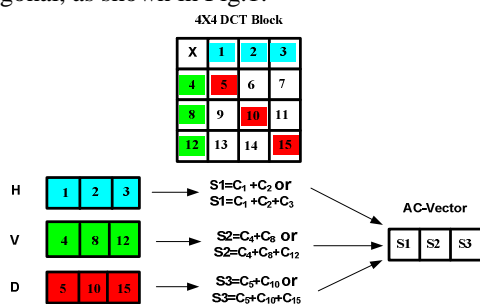

Figure 1: AC-Pattern

This selection is retained because of its ability represent local structure of content block. For efficiency we calculate the sum of 2 or 3 coefficients in each group and use these respective 3 summations to form the AC-Pattern. We use parameter Nc to represent the number of coefficients that we used in each group, thus Nc=2 or Nc=3. According to our experimental results, we use Nc=2. Thus we use only 6 coefficients to construct the AC-Pattern, and the dimension of this AC feature vector is 3. Compared with the method presented in (Zhong and Defée, 2005) (Bai et al., 2011), this selection can reduce the complexities of the feature vector obviously.

From the original histogram of AC-Pattern, we can make two observations: the first is that there is only part of AC-Patterns that appears in large quantities and a large number of AC-Patterns that appear rarely (Zhong and Defée, 2005). So in consideration of time-consuming and efficiency, we just select some of AC-Patterns which have higher frequency to construct the histogram. We use parameter ACbins to represent the number of AC-Patterns that are selected. For constructing the AC-Pattern histogram of an image, we just calculate the number of appearance of these AC-Patterns in this image, and then we get the AC-Pattern histogram H_{AC}. The second observation is that the first AC-Pattern inside the histogram is very dominant. This AC-Pattern mainly corresponds to blocks of image background and we will not consider this pattern in the AC-Pattern histogram by eliminating it.

So from these two observations the histogram of AC-Patterns that we will use for retrieval is as shown in Fig.2. In this histogram, we select first 70 (ACbins) high frequencies AC-Patterns.
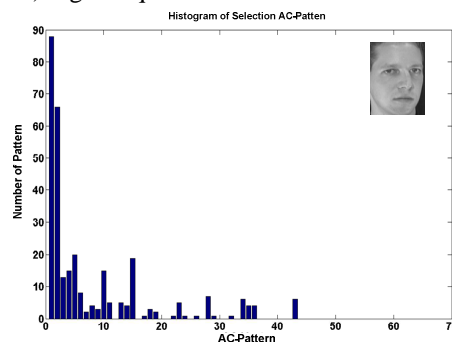

Figure 2: Histogram of selected AC-Pattern

2.4. DC-Pattern and its histogram

In complement to previous features that describes the local structures inside each block, we will observe, for each block and its neighbours, more global structures features by using gradients between blocks. To do so, DC-DirecVec (Zhong and Defée, 2005) is defined and used as feature for DC-Patterns. Like the same observations can be done in AC-

Pattern histogram, we select those dominant DC-Patterns to construct DC-Pattern histogram.

2.5. Feature vector descriptor and similarity measurement

We use the concatenation of AC-Pattern and DC-Pattern histogram to do image retrieval. In this context, the descriptors are defined as follows:

$$D = [(1-\alpha) \times H_{AC}, \alpha \times H_{DC}] \qquad (1)$$

α is a weight parameter that controls the impact of AC-Patterns and DC-Patterns histogram.

To measure the similarity between two descriptors we use the Manhattan distance:

$$Dis_{i,j} = \sum_{k=1}^{m} |D_i(k) - D_j(k)| \qquad (2)$$

where k demonstrates the components of the descriptor and i, j demonstrate the different descriptors, m indicates the total number of components in the descriptor, the dimension of descriptor, for example.

3 EXPERIMENTAL RESULTS

We perform retrieval experiments on ORL face database (AT&T Laboratories Cambridge) and also on VisTex texture database (Media Laboratory MIT).

3.1. Experiments on face database

The ORL database includes 10 different images of 40 peoples. For tests, we use first 6 images as image database and remaining 4 images as test images for retrieval. Therefore, the total number of images in the database is 240 and that of query images is 160.

For evaluating the performance we use Equal Error Rate (EER) (Bolle et al., 2000). If a value is used to express the similarity between query images and images in the database, so given a certain threshold, an input image of certain class A, may be recognized falsely as class B. Then the ratio of how many images of class A have been recognized as other class is called FRR (False Rejected Rate), while the ratio of how many images of other classes have been recognized into class A is call FAR (False Accept Rate). When both rates take equal values, an equal error rate (EER) is got. The lower the EER is, the better is the system's performance, as the total error rate is the sum of FAR and FRR.

We show the result of the experiment in which we use the concatenation of histogram of AC-Pattern and histogram of DC-Pattern to do image retrieval.

In all the experiments, the histogram of DC-Pattern is the same, but the approach to construct AC-Pattern is different. We name the method presented in (Zhong and Defée, 2005) as 'linear scan', the method presented in (Bai et al., 2011) as 'Adjacent zig-zag'. For both methods, we tested different sets of parameters to find the one that can assure the best performance. For our proposal, we use Nc=2, ACbins=70 and QPAC=30. And we adjust the weight parameter α to see the global comparison of the performance. The curves of the performance are as follows:
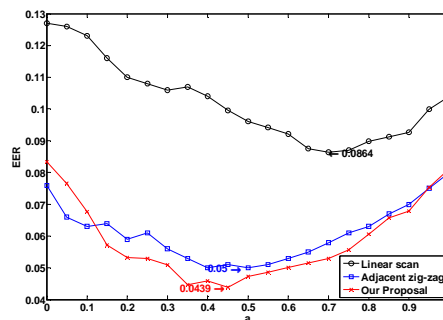

Figure 3: Performance of different method

It can be observed that the performance of retrieval can be improved by using the proposed approach to construct the descriptor.

We further compare the performance of the proposed method with those of other methods: Principal Component Analysis (Naz et al., 2006), 2D Principal Component Analysis (Xu et al., 2009) and Linear discriminate analysis (Goudelis et al., 2007). Table 1 shows the lowest EER reported in these methods and the lowest EER of our method. So we can conclude that our proposal outperforms all the other referenced methods.

Table 1: Comparison of EER with other methods

Method	PCA	2D PCA	LDA	Proposal
EER	0.095	0.165	0.11	0.0439

3.2. Experiments on texture database

As we want to extend our solution to a wider application field, we work with a selection of images from the popular MIT Vision Texture Database (VisTex), consisting of 40 textures which have already been extensively used in texture image retrieval literature (Do et al., 2002) (Kokare, 2005) (Huajing et al., 2008) (Kwitt et al., 2010). The 512x512 pixels colour versions of the textures are divided into 16 nonoverlapping subimages (128x128

pixels) and converted to gray scale images, thus creating a database of 640 images belonging to 40 class textures, each class with 16 different samples.

In retrieval experiments, a query image is each one of 640 images in our database. The relevant images for each query are all the subimages from the same original texture. We use the average retrieval rate (ARR) to evaluate the performance. For a given query image, the retrieval rate is defined as the percentage of the number of correct texture images retrieved in the same class as the query texture observed in the total number of retrieved images. For comparison purpose, we retrieve 16 images for each query. We use every subimage in the database as query to do retrieval and get the average retrieval rate finally.

Table 2 provides a quantized comparison. RCWF indicates Rotated Complex Wavelet Filters method proposed in (Kokare, 2005). CWT represents the Complex Wavelet Transform method presented in (Kingsbury et al., 1999). CWT+RCWF is also presented in (Kokare, 2005). PTR (Kwitt, 2010) is a probabilistic texture retrieval method based on dual-tree complex wavelet transform. It can be observed that our proposal outperforms other methods.

Table 2: Comparison of ARR with other methods

Method	RCWF	CWT	PTR	CWT+RCWF	Proposal
ARR(%)	75.78	80.78	81.73	82.34	83.64

4 CONCLUSIONS

In this paper we have presented a simple and effective approach for constructing descriptor using 2D DCT coefficients intended to image retrieval. Unlike other CBIR methods that usually focus on one kind of image database, our approach is suitable for different kind of image database. We evaluate our method both in widely used face database and texture database. From the point of view of recognition rate or average retrieval rate, the experimental results show higher performance compared to classical and state-of-art methods.

REFERENCES

AT&T Laboratories Cambridge, *ORL Database.* http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

Bai Cong, Kpalma Kidiyo and Ronsin Joseph, 2011. *Analysis of histogram descriptor for image retrieval in DCT domain.* Proceeding of the 4th International Conference on Intelligent Interactive Multimedia Systems and Services, Vol. 11, 227-235.

Bolle R.M.,Pankanti S. and Ratha N.K., 2000. *Evaluation techniques for biometrics-based authentication systems (FRR).* Proc. International Conf. on Pattern Recognition, Vol. 2, 831-837.

Do M. N. and Vetterli M., 2002. *Wavelet-based texture retrieval using generalized Guassian density and Kullback-leibler distance.* IEEE Trans. Image Process, Vol. 11, No. 2, 146–158.

Goudelis,G., Zafeiriou,S. and Tefas,A.,Pitas,I., 2007. *Class-Specific Kernel-Discriminant Analysis for Face Verification.* IEEE Transactions on Information Forensics and Security, Volume 2, 570 – 587.

Kim Yong-Ho, Lee Seok-Han and Lee Sangkeun, Kim Tae-Eun, Choi Jong-Soo, 2010. *A Novel Image Retrieval Scheme Using DCT Filter-Bank of Weighted Color Components.* 2010 International Conference on Information Science and Applications, 1-8.

Kokare M.,Biswas P.K. and Chatterji, B.N., 2005. *Texture image retrieval using new rotated complex wavelet filters.* IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 35, Issue: 6, 1168 - 1178.

Kwitt, R.; Uhl, A., 2010. *Lightweight Probabilistic Texture Retrieval.* IEEE Transactions on Image Processing, Vol.19, Issue: 1, 241 – 253.

Media Laboratory, MIT. *VisTex Database of textures.* http://vismod.media.mit.edu/vismod/imagery/VisionTexture/

Naz, E., Farooq, U., and Naz, T., 2006. *Analysis of Principal Component Analysis-Based and Fisher Discriminant Analysis-Based Face Recognition Algorithms.* 2006 International Conference on Emerging Technologies, 121 – 127.

Tsai Tienwei, Huang Yo-Ping, and Chiang Tei-Wei, 2006. *Image Retrieval Based on Dominant Texture Features.* 2006 IEEE International Symposium on Industrial Electronics, 441-446.

Xu Zhijie, Zhang Jianqin, and Dai Xiwu, 2009. *Boosting for Learning a Similarity Measure in 2DPCA Based Face Recognition.* 2009 World Congress on Computer Science and Information Engineering, Vol. 7, 130 – 134.

Zhong D., Defée I. 2005. *DCT histogram optimization for image database retrieval.* Pattern Recognition Letters, Vol. 26, 2272-2281.